10/727,010

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 9/00 — A2

(11) International Publication Number: **WO 99/53037**

(43) International Publication Date: 21 October 1999 (21.10.99)

(21) International Application Number: PCT/US99/08202

(22) International Filing Date: 14 April 1999 (14.04.99)

(30) Priority Data:
60/081,846 15 April 1998 (15.04.98) US

(71) Applicant *(for all designated States except US)*: PRODIGENE, INC. [US/US]; Suite 220, 1500 Research Parkway, College Station, TX 77845 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: JILKA, Joseph, M. [US/US]; Suite 220, 1500 Research Parkway, College Station, TX 77845 (US).

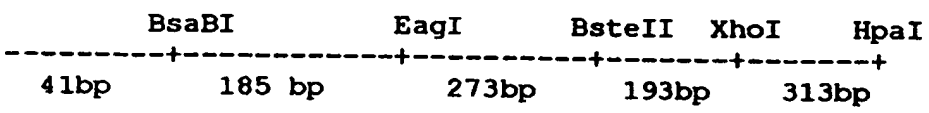
(74) Agent: NEBEL, Heidi, S.; Zarley, McKee, Thomte, Voorhees & Sease, Suite 3200, 801 Grand Avenue, Des Moines, IA 50309-2721 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: OPTIMIZED NUCLEOTIDE SEQUENCE ENCODING ORGANOPHOSPHOROUS HYDROLASE AND METHODS OF USE FOR SAME



(57) Abstract

A nucleotide sequence is disclosed which is optimized for plant expression and which encodes organophosphorous hydrolase. Expression constructs, transformation vectors as well as transformed cells are also disclosed which achieve expression levels superior to those demonstrated by expression in bacteria such as *E. coli* for high throughput production of the enzyme. Transgenic plants may themselves be used or the protein may be harvested from said plants for *in situ* or other environmental detoxification of organophosphorous neurotoxin contaminated areas.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE: OPTIMIZED NUCLEOTIDE SEQUENCE ENCODING ORGANOPHOSPHOROUS HYDROLASE AND METHODS OF USE FOR SAME

CROSS REFERENCE TO RELATED APPLICATION

This is a continuation application of co-pending provisional application, Serial No. 60/081,846 filed April 15, 1998.

BACKGROUND OF THE INVENTION

Synthetic organophosphorous neurotoxins are used extensively as agricultural and domestic pesticides, including insecticides, fungicides, and herbicides. Naturally occurring bacterial isolates capable of metabolizing this class of compounds have received considerable attention since they provide the possibility of both environmental and *in situ* detoxification. *Pseudomonas putida* MG, *Pseudomonas diminuta, and Flavobacterium*, s.pp. have been shown to possess the ability to degrade an extremely broad spectrum of organophosphorous phosphotriesters as well as thiol esters. (McDaniel, et al., "Cloning and Sequencing of a Plasmid-Borne Gene (opd) Encoding a Phosphotriesterase," Journal of Bacteriology, Vol. 170:5, 2306-11 (May 1988).)

Organophosphorous hydrolase (OPH) is a broad spectrum OP hydrolase that is capable of detoxifying organophosphorous neurotoxins by creating various phosphoryl bonds (P-O, P-F, P-CN, and P-S) between the phosphorous center and varying electrophilic leaving groups. (Dave et al., "Characterization of Organophosphorous Hydrolases and the Genetic Manipulation of the Phosphotriesterase from *Pseudomonas diminuta*", Chemical-Biological Interactions 87, 55-68 (1993)). This enzyme is often identified by more limited descriptors such as phosphotriesterase, DFPase, parathion hydrolase, parathion aryl esterase, or paraoxonase as well as being called a somanase or sarinase. The hydrolytic reaction rates with several phosphotriesterases appears to be limited by diffusion to the active center of the enzyme. (Caldwell et al., "Limits of Diffusion in the Hydrolysis of

Substrates of the Phosphotriesterase From *Pseudomonas diminuta*", Biochemistry 30, 7438-7444 (1991)).

This broad spectrum hydrolase (OPH) is the only enzyme which has been shown to be able to hydrolyze the P-S bond of various phosphorothioate pesticides. The toxicity of hydrolyzed products has been shown to be significantly reduced as indicated by the loss of inhibition of acetyl cholinesterase activity and by decreased neurotoxic response in animals. (Kolakowski, et al., Biocatolisis and Biotransformation, Vol. 15, 297-312 (1997)). OP-thioate insecticides (acephate, azinophos-ethyl, demeton-S, malathion and phosalone) have been shown to be hydrolyzed by OPH. The hydrolysis of these pesticides has a first order dependency on the amount of enzyme used and the reaction time. The enzyme hydrolyzed acephate, azinophos-ethyl, demeton-S and phosalone at relatively fast velocities with reaction rates which are thousands of times greater than that which occurs during strong alkaline hydrolysis. In contrast, the enzyme possessed poor capability for a malathion hydrolysis, although still significantly better than non-enzymatic hydrolysis under similar conditions. When compared to the hydrolysis of P-O bond phosphotriester substrates and P-F bond phosphofluoridate substrates, the thioesters (P-S bond esters) hydrolysis was much slower in general. (See Kolakowski, et al., Biokatolisis and Biotransformation, Vol., 297-312 (1997)).

The genetic expression systems in the purification procedures of recombinant OPH from *E. coli* have been described previously in Lai et al., 1994. However, the expression in *E. coli* is approximately 5-10 milligrams/liter. This level is inadequate for the intended uses such as detoxification of pesticides on a commercial level. OPH has also been expressed in baculovirus systems with similar expression levels as those in *E. coli*. As can be seen a need exists in the art for a reliable high expression system for production of OPH.

The present invention relates to the plant expression of OPH as well as the creation of a synthetic OPH nucleotide sequence which can be expressed in

plant systems or yeast systems utilizing appropriate leader sequences whereby the expression level of OPH is greatly increased over that of *E. coli* or baculovirus systems. The present invention relates not only to the creation of the optimized OPH coding sequence, but also to transformation, expression of the same, and useof the recombinant protein in the hydrolysis of a variety of organophosphorous neurotoxins such as those in many widely used agricultural and domestic pesticides.

SUMMARY OF THE INVENTION

The present invention is directed to a nucleotide sequence which encodes OPH and its expression in plant systems. In addition, the present invention is directed to a nucleotide sequence which encodes an optimized OPH gene preferrably with a host specific leader sequence. The present invention is further directed to expression constructs, vectors and transgenic plants which contain these OPH nucleotide sequences. In yet another embodiment the invention is directed to the expression of OPH in plant or yeast cells. The present invention is also directed to methods for using transgenic plants which express this OPH to detoxify environments which are contaminated with insecticides and other neurotoxins. More specifically, the present invention is directed to transgenic plants which are capable of expressing OPH at levels greater than that of the typical 10-15 mg/l of *E. coli* or baculovirus.

It is to be understood that the OPH nucleotide sequence of the present invention can be utilized in conjunction with a wide variety of host organisms including plants, bacteria and yeast. It should also be understood that trivial modifications to the OPH gene of this invention are encompassed in this invention. Further, while it is to be understood that the expression of the OPH of the present invention can be utilized either as present in the transgenic plants or harvested therefrom to hydrolyze neurotoxins such as pesticides, this expressed OPH gene can further be utilized to hydrolyze other chemical agents or compounds containing P-O, P-S, P-CN, and P-S bonds.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 is a schematic of optimized OPH nucleotide sequence showing the restriction sites used in the synthesis of the gene.

Fig. 2 is an optimized nucleotide sequence encoding OPH gene (SEQ ID NO:3).

Fig. 3A and 3B are a nucleotide sequence which encodes plant leader sequence (SEQ ID NO:7).

Fig. 4A and 4B are a nucleotide sequence which encodes OPH and which also contains a plant leader sequence (SEQ ID NO:5).

DETAILED DESCRIPTION OF THE INVENTION

The present invention includes novel nucleotide sequences which are optimized to encode an organophosphorous hydrolase protein. An expression construct comprising this sequence and regulatory elements for OPH expression in a host organism is also provided. The present invention also includes plants, seeds, and plant tissue capable of expressing the novel nucleotide sequence.

In order to provide a clear and consistent understanding of the specification and the claims, including the scope given to such terms, the following definitions are provided:

Coding DNA Sequence: A DNA sequence from which the information for making a peptide molecule, mRNA, or tRNA are transcribed. A DNA sequence may be a gene, combination of genes, or a gene fragment.

Gene: A chromosomal region which is responsible for a cellular product.

Microorganism: A member of one of the following classes; bacteria, fungi, protozoa or viruses.

Plant Tissue: Any tissue of a plant in plant or in culture. This term includes, but is not limited to, whole plants, plant cells, plant organs, plant seeds, embryos, pollen, silk, tassel, leaf, root, protoplasts, callus, cell cultures and any other group of plant cells organized into structural and/or functional

units. The use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

Plant Transformation Vector: A plant, bacterial or viral vector that is capable of transforming plant tissue such that following transformation the plant tissue contains the DNA to be introduced by the vector in the plant tissues.

Substantial Sequence Homology: Substantial functional and/or structural equivalents between sequences of nucleotides or amino acids. Functional and/or Structural Differences between sequences having substantial sequence homology will be diminimous.

Synthetic Gene: A DNA sequence that exists in total or in part through manufacture by manufacture *in vitro*.

Transgenic Plant: a plant which contains and expresses DNA that was not pre-existing in the plant, either at all or in the post-transformation quantity, prior to the introduction of the DNA into the plant.

Organophosphorous hydrolase protein: as used herein this term shall mean any protein or protein fragment which maintains the functional properties of OPH as determined by the assays described in Daie et al., "Characterization of Organophosphorous Hydrolases and the Genetic Manipulation of the Phosphotriesterase from *Pseudomonas diminuta*", Chemical-Biological Interactions, 87, 55-68 (1993), and Caldwell et al., "Limits of Diffusion in the Hydrolysis of Substrates of the Phosphotriesterase From *Pseudomonas diminuta*", Biochemistry 30, 7438-7444 (1991). Typically this will be an amino acid with at least 30%-50% homology with at least one form of the protein as disclosed herein. 80% homology is preferred and 90% homology is most preferred especially including conservative substitutions. Homology is calculated by standard methods which involve aligning two sequences to be compared so that the maximum matching occurs, and calculating the percentage of matches. Substantially equivalent substances to these include those wherein one or more of the residues of the native sequence

is deleted, substituted for, or inserted by a different amino acid or acids. Preferred substitutions are those which are conservative, i.e., wherein a residue is replaced by another of the same general type. As is well understood, naturally occurring amino acids can be sub classified as acidic, basic, neutral and polar, or neutral and nonpolar. Furthermore, three of the encoded amino acids are aromatic. It is generally preferred that peptides differing from the native OPH sequence contain substitutions which are from the same group as that of the amio acid replaced. Thus, in general, the basic amino acids Lys and Arg are interchangeable; the acidic amino acids aspartic and glutamic are interchangeable; the neutral polar amino acids Ser, Thr, Cys, Gln, and Asn are interchangeable; the nonpolar aliphatic acids Gly, Ala, Val, Ile, and Leu are conservative with respect to each other (but because of size, Gly and Ala are more closely related and Val, Ile and Leu are more closely related), and the aromatic amino acids Phe, Trp, and Tyr are interchangeable. While proline is a non-olar neutral amino acid, it represents difficulties because of its effects on conformation, and substitutions by or for proline are not preferred, except when the same or similar conformational results can be obtained. Polar amino acids which represent conservative changes include Ser, Thr, Gln, Asn; and to a lesser extent, Met. In addition, although classified in different categories, Ala, Gly, and Ser seem to be interchangeable, and Cys additionally fits into this group, or may be classified with the polar neutral amino acids. In general, whatever substitutions are made are such that the functional properties of the intact proteinaceous molecule is retained and ancillary properties, such as non-toxicity are not substantially disturbed as described earlier.

Conditions of high stringency: as used herein this term shall mean nucleotide hybridization conditions equivalent to the following: 42°C in a buffer containing 50% formamide, 1M sodium chloride, 1% SDS, 10% dextran sulfate, 100 μg/ml denatured salmon sperm, or those described in Reppert, S.M. et al. (1994) Neuron 13:1177-1185.

Optimized OPH Gene

The amino acid and nucleic acid sequence for the native plasmid-borne organophosphorus degrading from Pseudomonas diminuta has been determined. McDaniel et al., "Cloning and Sequencing of a Plasmid-Born Gene (opd) Encoding a Phosphotriesterase," Journal of Bacteriology, 170 (5):230-11, (May 1988). The amino acid sequence was back-translated into a DNA coding sequence using a maize codon preference table (http://www.gcg.com/techsupport/data/maize-high.cod). This sequence, which was optimized for corn, was searched for putative deleterious sequences compiled in Table 1 of Example 1. Codons were chosen so that deleterious mRNA signals would be eliminated in the DNA coding sequence. In addition, convenient restriction enzyme cites were added to facilitate downstream cloning. The optimized DNA coding sequence was divided into convenient cloning lengths and then each of these lengths were further subdivided into oligonucleotides of approximately 50 bases in length with 10 base overhangs. Oligonucleotides corresponding to both the sense and antisense strand were synthesized and 5' phosphorylated. The oligonucleotides were then annealed and ligated. Following ligation, varying dilutions were subjected to PCR using primers specific to the 5' and 3' ends. PCR product was analyzed using polyacrylamide gels, and DNA of the proper expected lengths was extracted from the gel and the DNA eluted from the extracted gel material. The DNA was cloned into a pCR 2.1 TOPO(Invitrogen) plasmid as per the manufacturer's instructions. Transformed colonies containing the PCR insert were grown overnight and the plasmid DNA purified. The plasmid DNA was then analyzed for the correct PCR insert as per standard protocols known to those of skill in the art. The inserts were sequenced to verify that the correct sequence was created and the optimized gene was assembled properly using standard protocols known to those of skill in the art.

Appropriate Leader Sequence

To enable the OPH gene to be expressed in a host organism, an appropriate leader sequence or subcellular localization signal is preferably

integrated onto the OPH encoding nucleotide sequence gene. For example, a plant leader sequence will ensure high expression in plants while a yeast leader sequence will express high levels of OPH expression in yeast. Upon assembly of the OPH gene with a 5' leader sequence specific for the host organism of choice, the OPH gene containing the appropriate leader can be cloned into the standard expression cassette for that host to ensure high expression in that organism. For example, a 5' fragment can be assembled to contain a plant leader sequence to ensure high expression in plants. The OPH gene containing the plant leader may be cloned into a standard plant upon expression cassette to ensure high expression in plants.

Transformation techniques

In the early 1980s, it became clear that plants were amenable to genetic engineering because foreign genes can be stably introduced into plant chromosomal DNA by a variety of techniques. (Gasser, C. S. & Fraley, R. T., "Genetically Engineering Plants for Crop Improvement," Science, 224, 1293 (1989); Potrykus, I., "Gene Transfer to Plants; Assessment of Published Approaches and Results," Annu. Rev. Plant Physiol. Plant Mol. Biol., 42, 205 (1991).) If a given foreign gene contains the appropriate regulatory sequences, the gene product will be synthesized by the transformed plant. Furthermore, most plant cells are totipotent, thereby allowing regeneration of a fertile "transgenic" plant from a single transformed cell. If the transgenic plant flowers and produces viable seed, the acquired trait will be preserved in the progeny. (Greenberg & Glick, "The Use of Recombinet DNA Technology to Produce Genetically Modified Plants," Methods in Plant Molecular Biology and Biotechnology (Glick & Thompson, eds.) (CRC Press 1993).

Once a nucleotide acid sequence of interest is isolated or synthesized, it is included into an expression cassette which typically comprises a transcription unit of a promoter operably linked to the nucleotide sequence which is functional in a plant cell and a termination or polyadenylation signal. This expression cassette typically forms part of a plant transformation vector. This plant transformation vector will also contain the appropriate regulatory

sequences in addition to the gene of interest. The vector DNA facilitates manipulation of the gene in the host cell, such as *E. coli.* or yeast, prior to plant transformation. In addition, the host vector-cell can act as a vehicle to transfer the gene to the host plant such as the case of the plant pathogenic bacterium, *Agrobacterium tumefaciens.* An idealized vector would contain a multiple cloning site, an antibiotic resistance or other selective gene, a broad-host bacterial origin of replication, and a selection marker or gene for selection of the foreign DNA in transformed plants.

PROMOTERS

The constructs, promoters or control systems used in the methods of the invention may include a tissue specific promoter, an inducible promoter or a constitutive promoter.

A large number of suitable promoter systems are available. For example one constitutive promoter useful for the invention is the cauliflower mosaic virus (CaMV) 35S. It has been shown to be highly active in many plant organs and during many stages of development when integrated into the genome of transgenic plants including tobacco and petunia, and has been shown to confer expression in protoplasts of both dicots and monocots.

Organ-specific promoters are also well known. For example, the E8 promoter is only transcriptionally activated during tomato fruit ripening, and can be used to target gene expression in ripening tomato fruit (Deikman and Fischer, EMBO J. (1988) 7:3315; Giovannoni et al., The Plant Cell (1989) 1:53). The activity of the E8 promoter is not limited to tomato fruit, but is thought to be compatible with any system wherein ethylene activates biological processes. Similarly the Lipoxegenase ("the LOX gene") is a fruit specific promoter.

Seed specific promoters include the Napin promoter described in united States Patent 5,110,728 to Calgene, which describes and discloses the use of the napin promoter in directing the expression to seed tissue of an acyl carrier protein to enhance seed oil production; the DC3 promoter from carrots which is early to mid embryo specific and is disclosed at Plant Physiology, Oct. 1992

100(2) p. 576-581, "Hormonal and Environmental Regulation of the Carrot Lea-class Gene Dc 3, and Plant Mol. Biol., April 1992, 18(6) p. 1049-1063, "Transcriptional Regulation of a Seed Specific Carrot Gene, DC 8": the phaseolin promoter described in United States Patent 5,504,200 to Mycogen which discloses the gene sequence and regulatory regions for phaseolin, a protein isolated from *P. vulgaris* which is expressed only while the seed is developing within the pod, and only in tissues involved in seed generation.

Other organ-specific promoters appropriate for a desired target organ can be isolated using known procedures. These control sequences are generally associated with genes uniquely expressed in the desired organ. In a typical higher plant, each organ has thousands of mRNAs that are absent from other organ systems (reviewed in Goldberg, Phil, Trans. R. Soc. London (1986) B314-343. mRNAs are first isolated to obtain suitable probes for retrieval of the appropriate genomic sequence which retains the presence of the natively associated control sequences. An example of the use of techniques to obtain the cDNA associated with mRNA specific to avocado fruit is found in Christoffersen et al., Plant Molecular Biology (1984) 3:385. Briefly, mRNA was isolated from ripening avocado fruit and used to make a cDNA library. Clones in the library were identified that hybridized with labeled RNA isolated from ripening avocado fruit, but that did not hybridize with labeled RNAs isolated from unripe avocado fruit. Many of these clones represent mRNAs encoded by genes that are transcriptionally activated at the onset of avocado fruit ripening.

Another very important method that can be used to identify cell type specific promoters that allow even to identification of genes expressed in a single cell is enhancer detection (O'Kane, C., and Gehring, W.J. (1987), "Detection *in situ* of genomic regulatory elements in *Drosophila*", Proc. Natl. Acad. Sci. USA, 84, 9123-9127). This method was first developed in *Drosophila* and rapidly adapted to mice and plants (Wilson, C., Pearson, R.K., Bellen, H.J., O'Kane, C.J., Grossniklaus, U., and Gehring, W.J. (1989), "P-element-mediated enhancer detection: an efficient method for isolating and

characterizing developmentally regulated genes in *Drosophila*", Genes & Dev., 3, 1301-1313; Skarnes, W.C. (1990), "Entrapment vectors: a new tool for mammalian genetics", Biotechnology, 8, 827-831; Topping, J.F., Wei, W., and Lindsey, K. (1991), "Functional tagging of regulatory elements in the plant genome", Development, 112, 1009-1019; Sundaresan, V., Springer, P.S., Volpe, T., Haward, S., Jones, J.D.G., Dean, C., Ma, H., and Martienssen, R.A., (1995), "Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements", Genes & Dev., 9, 1797-1810).

The promoter used in the method of the invention may also be an inducible promoter. An inducible promoter is a promoter that is capable of directly or indirectly activating transcription of a DNA sequence in response to an inducer. In the absence of an inducer, the DNA sequence will not be transcribed. Typically, the protein factor that binds specifically to an inducible promoter to activate transcription is present in an inactive form which is then directly or indirectly converted to the active form by the inducer. The inducer may be a chemical agent such as a protein, metabolite (sugar, alcohol etc.), a growth regulator, herbicide, or a phenolic compound or a physiological stress imposed directly by heat, salt, toxic elements etc. or indirectly through the action of a pathogen or disease agent such as a virus. A plant cell containing an inducible promoter may be exposed to an inducer by externally applying the inducer to the cell such as by spraying, watering, heating, or similar methods. Examples of inducible promoters include the inducible 70 kd heat shock promoter of D. melanogaster (Freeling, M., Bennet, D.C., Maize ADN 1, Ann. Rev. of Genetics, 19:297-323) and the alcohol dehydrogenase promoter which is induced by ethanol (Nagao, R.T., et al., Miflin, B.J., Ed. Oxford Surveys of Plant Molecular and Cell Biology, Vol. 3, p. 384-438, Oxford University Press, Oxford 1986) or the Lex A promoter which is triggered with chemical treatment and is available through Ligand pharmaceuticals. The inducible promoter may be in an induced state throughout seed formation or at least for a period which corresponds to the transcription of the DNA sequence of the recombinant DNA molecule(s).

Another example of an inducible promoter is the chemically inducible gene promoter sequence isolated from a 27 kd subunit of the maize glutathione-S-transferase (GST II) gene. Two of the inducers for this promoter are N,N-diallyl-2,2-dichloroacetamide (common name: dichloramid) or benzyl-=2-chloro-4-(trifluoromethyl)-5-thiazolecarboxylate (common name: flurazole). In addition, a number of other potential inducers may be used with this promoter as described in published PCT Application No. PCT/GB90/00110 by ICI.

Another example of an inducible promoter is the light inducible chlorophyll a/b binding protein (CAB) promoter, also described in published PCT Application No. PCT/GB90/00110 by ICI.

Inducible promoters have also been described in published Application No. EP89/103888.7 by Ciba-Geigy. In this application, a number of inducible promoters are identified, including the PR protein genes, especially the tobacco PR protein genes, such as PR-1a, PR-1b, PR-1c, PR-1, PR-A, PR-S, the cucumber chitinase gene, and the acidic and basic tobacco beta-1,3-glucanase genes. There are numerous potential inducers for these promoters, as described in Application No. EP89/103888.7.

The preferred promoters may be used in conjunction with naturally occurring flanking coding or transcribed sequences of the seed specific Polycomb genes or with any other coding or transcribed sequence that is critical to Polycomb formation and/or function.

It may also be desirable to include some intron sequences in the promoter constructs since the inclusion of intron sequences in the coding region may result in enhanced expression and specificity. Thus, it may be advantageous to join the DNA sequences to be expressed to a promoter sequence that contains the first intron and exon sequences of a polypeptide which is unique to cells/tissues of a plant critical to seed specific Polycomb formation and/or function.

Additionally, regions of one promoter may be joined to regions from a different promoter in order to obtain the desired promoter activity resulting in

a chimeric promoter. Synthetic promoters which regulate gene expression may also be used. The expression system may be further optimized by employing supplemental elements such as transcription terminators and/or enhancer elements.

OTHER REGULATORY ELEMENTS

In addition to a promoter sequence, an expression cassette or construct should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region or polyadenylation signal may be obtained from the same gene as the promoter sequence or may be obtained from different genes. Polyadenylation sequences include, but are not limited to the *Agrobacterium* octopine synthase signal (Gielen et al., EMBO J. (1984) 3:835-846) or the nopaline synthase signal (Depicker et al., Mol. and Appl. Genet. (1982) 1:561-573).

MARKER GENES

Recombinant DNA molecules containing any of the DNA sequences and promoters described herein may additionally contain selection marker genes which encode a selection gene product which confer on a plant cell resistance to a chemical agent or physiological stress, or confers a distinguishable phenotypic characteristic to the cells such that plant cells transformed with the recombinant DNA molecule may be easily selected using a selective agent. One such selection marker gene is neomycin phosphotransferase (NPT II) which confers resistance to kanamycin and the antibiotic G-418. Cells transformed with this selection marker gene may be selected for by assaying for the presence in vitro of phosphorylation of kanamycin using techniques described in the literature or by testing for the presence of the mRNA coding for the NPT II gene by Northern blot analysis in RNA from the tissue of the transformed plant. Polymerase chain reactions are also used to identify the presence of a transgene or expression using reverse transcriptase PCR amplification to monitor expression and PCR on genomic DNA. Other commonly used selection markers include the ampicillin resistance gene, the tetracycline resistance and the hygromycin resistance gene. Transformed

plant cells thus selected can be induced to differentiate into plant structures which will eventually yield whole plants. It is to be understood that a selection marker gene may also be native to a plant.

TRANSFORMATION

Various methods are known in the art to accomplish the genetic transformation of plants and plant tissues. Some of these methods include transformation, or agroinfection, by *Agrobacterium tumefaciens*, electroporation, particle bombardment or biollistic projection, or direct gene transfer.

Various methods are known in the art to accomplish the genetic transformation of plants and plant tissues. Among these methods for introducing foreign DNA into plants is *Agrobacterium* species transformation and transformation by direct gene transfer.

Agrobacterium tumefaciens is the etiologic agent of crown gall. The wild type form of *Agrobacterium tumefaciens* carries the Ti (tumor-inducing) plasmid that directs the production of tumorigenic crown gall growth on the host plants. The crown gall is produced following the transfer of the tumor inducing T-DNA region from the Ti plasmid into the genome of an infected plant. This DNA fragment encodes genes for auxin and cytokinin biosynthesis, and it is these hormones in high concentration that promote growth of undifferentiated cells in the crown gall. Transfer of the T-DNA to the plant genome requires that the Ti plasmid-encoded virulence genes as well as the T-DNA borders, a set of direct DNA repeats that delineate the region to be transferred. The tumor inducing genes can be removed from Ti plasmid vectors, disarming the pathogenic nature of the system, without affecting the transfer of DNA fragments between the T-DNA borders. Therefore, the tumor inducing genes are generally replaced with a gene encoding resistance to kanamycin, or some other gene, to allow for selection of transformants, and a gene encoding the desired trait. The *Agrobacterium* containing the engineered plasmid is co-cultivated with cultured plant cells or wounded tissue. The de-differentiated plant cells are then propagated on selective media, and a

transgenic plant is subsequently regenerated from the transformed cells by altering the levels of auxin and cytokinin in the growth medium.

Current protocols for *Agrobacterium* mediated transformation often employ binary vector systems, which divide the Ti plasmid into two components, a shuttle vector and a helper plasmid. The helper plasmid, which is permanently placed in the *Agrobacterium* host, carries the virulence genes. However, a much smaller shuttle vector contains T-DNA borders, a broad-host range bacterial origin of replication, antibiotic resistance markers, and a multiple cloning site for incorporation of the foreign gene. In the alternative, a similar strategy employs cointegrating Ti plasmid vectors, whereby an intermediate plasmid containing antibiotic resistance, the gene to be transferred and one T-DNA border are used to transform *A. tumefaciens* containing a disarmed Ti plasmid possessing the virulence genes and one T-DNA border. The two plasmids homologously recombined in vivo at the T-DNA borders placing the antibiotic resistance gene and the gene of interest between two T-DNA borders, one from each plasmid. The genes are then transferred into plant tissue upon co-cultivation.

The *Agrobacterium* system has been well studied and has further been developed into a system which permits routine transformation of a variety of plant tissues. (See, e.g., Schell, J. et al., Bio/Technology, 1:175 (1983); Chilton, M-D, Scientific American, 248:50 (1983)). Some of the tissues transformed utilizing *Agrobacterium* include tobacco, Barton, K.A. et al., Cell, 32:1033 (1983); tomato, Fillatti, J. et al., Bio/Technology, 5:726 (1987); sunflower, Everett, N. P. et al., Bio/Technology, 5:1201 (1987); cotton, Umbeck, P. et al., Bio/Technology, 5:263 (1987); canola, Pua, E. C. et al., Bio/Technology, 5:815 (1987); potato, Facciotti, D. et al., Bio/Technology, 3:241 (1985); poplar, Pythoud, F. et al., Bio/Technology, 5:1323 (1987); and soybean, Hinchee, M. A. et al., Bio/Technology, 6:915 (1988).

Another vector for biological plant transformation is *Agrobacterium rhizogenes*. *A. rhizogenes*, which incite root formation in many dicotyledonous plant species, carries the Ri (root-inducing) plasmid which functions in a

manner analogous to the Ti plasmid of *A. tumefaciens*. Transformation using *A. rhizogenes* has also been successfully utilized to transform plants, for example, alfalfa, Sukhapinda, K. et al., Plant Mol. Biol., 8:209 (1987); *Solanum nigrum* L., Wei Z-H. et al., Plant Cell Reports 5:93 (1986); and, poplar, Pythoud, et al., supra.

Several direct gene transfer procedures have also been developed to transform plants and plant tissues. In the direct transformation of protoplasts, the uptake of exogenous genetic material into a protoplast may be enhanced by use of a chemical agent or electric field. The exogenous material may then be integrated into the nuclear genome. Microprojectile bombardment, electroporation in addition to several other direct transformation methods exist and are known to those of skill in the art.

Electroporation is another effective means of introducing foreign DNA into plant cells. (Saul, et al., "Direct DNA Transfer to Protoplasts With and Without Electroporation," Plant Molecular Biology Manual, Vol. A1, Kluwer Academic Publishers, Dordrecht, 1988, 1.) This technique involves transferring naked DNA into cells utilizing electrophoretic means. Since any DNA fragment can be delivered to the cell, this technique has the advantage of allowing assimilation of a gene without having to clone the DNA into a host vector such as *A. tumefaciens*.

Another technique for delivering DNA to intact plant tissue or protoplasts is biollistic projection or microprojectile bombardment. (Tomes, et al., "Transgenic Tobacco Plants and their Progeny Derived by Microprojectile Bombardment of Tobacco Leaves," Plant Mol. Biol., 14, 21 (1990); SVAB, et al., "Stable Transformation of Plastids of Higher Plants," Proc. Nat'l. Acad. Sci. U.S.A., 87, 8526 (1990).) With this technique, microprojectile particles (e.g., 1.2 micrometer gold or Tungsten beads) coated with DNA are accelerated at high speeds into plant tissue. This approach has the distinct advantage of being applicable to any intact plant tissue or region of the plant and has been used to transform organellar DNA. (SVAB, et al. (1990); Boynton, et al., "Chloroplast Transformation in *Chlamydomonas* with High Velocity

Microprojectiles," Science, 240, 534 (1988).) This technique is especially useful for plants that are otherwise recalcitrant with respect to transformation and/or regeneration. For instance, biollistic methods have been successfully used to produce transgenic monocot cereal plants. (Potrykus, I., at Annu. Rev. Plant Physiol. Plant Mol. Biol., 42, 205 (1991); Wu, et al., "Transformation and Regeneration of Important Crop Plants; Rice is the Model System for Monocots," in Gene Manipulation and Plant Improvement, Vol. 2 (Gustafson, J. P., ed.) (Quantum Press, New York 1990) 251; Gordon-Kamm, et al., "Transformation of Maize Cells and Regeneration of Fertile Transgenic Plants," Plant Cell, 2, 603 (1990).)

DNA viruses have been used as gene vectors. Both Cauliflower Mosaic Virus (CaMV) and the closely related Figwort Virus are circular duplex DNA viruses which replicate via transcription of a full-length (35S) genomic RNA intermediate. A cauliflower mosaic virus carrying a modified bacterial methotrexate-resistance gene was used to infect a plant, whereby the foreign gene was systemically spread in the plant. (Brisson, N. et al., "Expression of a Bacterial Gene in Plants by Using a Viral Vector," Nature, 310:511 (1984); Brisson, et al. Methods for Plant Molecular Biology, p. 437 (1988).) The strong promoter responsible for the genomic replication of the CaMV virus (the 35S promoter) has been extensively exploited for the expression of heterologous genes in plants. Other advantages of this system are the ease of infection, systemically spread within the plant, and multiple copies of the gene per cell.

Exogenous DNA can be introduced into cells or protoplasts by microinjection. Microinjection is a method where a solution of plasmid DNA is injected directly into the cell with a finely pulled glass needle. Alfalfa protoplasts have been transformed utilizing this method with a variety of plasmids. Reich, T.J. et al., Bio/Technology 4:1001 (1986).

In liposome fusion, protoplasts and liposomees carrying the foreign gene of interest are brought together. As membranes merge, the foreign gene is transferred to the protoplast. Deshayes, A. et al., "Liposome-Mediated

Transformation of Tobacco Mesophyl Protoplasts by an *Escherichia coli* Plasmid," EMBO J, 4:2731 (1985).

A form of chemical mediated transformation utilizes polyethylene glycol (PEG) and has been carried out in *N. tabacum* a dicot, and *Lolium multiflorum*, a monocot. It is a chemical procedure of direct gene transfer based on synergistic interaction between Mg^{2+}, PEG, and possibly Ca^{2+}. Negrutiu, R. et al., Plant Mol. Biol., 8:363 (1987).

There are a variety of methods known to those of skill in the art for the regeneration of plants from plant tissue. See e.g., The Maize Handbook, (Freeling & Walbot, eds.) (Springer-Verlag, pubs. 1994) at pages 263-76. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated. It is possible to regenerate many species of plants from callus tissue. Both monocot plants, such as corn, rice, barley, wheat and rye, and dicot plants, such as sunflower, soybean, cotton, canola and tobacco, can be generated from callus.

In addition, plants can be regenerated from tissue transformed with *A. tumefaciens*. These include sunflower, Everatt, N.P. et al., supra; tomato, Fillatti, J.J. et al., supra; white clover, White, D.W. R. et al., Plant Mol. biol. 8:461 (1987); canola (rapeseed), Pua, E-C, et al., supra; cotton, Umbeck, P. et al., supra; tobacco, Horsch, R.B. et al., Science 225:1229 (1985) and Hererra-Estrella, L. et al., Nature, 303:209 (1983); and poplar, Pythoud et al., supra. The regeneration of alfalfa from tissue transformed with *A. rhizogenes* has been demonstrated by Sukhapinda, K., et al., supra.

Plant regeneration from protoplasts is a particularly useful technique. See Evans, D.A. et al., Handbook of Plant Cell Culture, 1:124 (1983). When a plant species can be regenerated from protoplasts, then direct gene transfer procedures can be utilized, and transformation is not dependent on the use of *A. tumefaciens*. Regeneration of plants from protoplasts has been demonstrated for rice, Abdullah, R. et al., Bio/Technology 4:1087 (1987); tobacco, Potrykus, I. et al., supra; canola, Kansha, et al., Plant Cell Reports 5:101 (1986); potato, Tavazza, R. et al., Plant Cell Reports 5:243 (1986);

eggplant, Sihackaki, D. et al., Plant Cell, Tissue, Organ Culture 11:179 (1987); cucumber, Jia S-R., et al., J. Plant Physiol. 124:393 (1986); poplar, Russel, J.A. et al., Plant Sci. 46:133 (1986); corn, Rhodes, C. et al., supra; and soybean, McCabe, D.E. et al., supra.

The organohydrolase may be harvested from the plant or the plant itself may be used to detoxify sites contaminated by organophosphorous neurotoxins, from insecticides, pesticides or even for use in chemical warfare detoxification.

After the expression cassette is stably incorporated into regenerated transgenic plants, it can be transferred to other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

It may be useful to generate a number of individual transformed plants with any recombinant construct in order to recover plants free from any position effects. It may also be preferable to select plants that contain more than one copy of the introduced recombinant DNA molecule such that high levels of expression of the recombinant molecule are obtained.

As indicated above, it may be desirable to produce plant lines which are homozygous for a particular gene. In some species this is accomplished rather easily by the use of anther culture or isolated microspore culture. This is especially true for the oil seed crop *Brassica napus* (Keller and Armstrong, Z. flanzenzucht 80:100-108, 1978). By using these techniques, it is possible to produce a haploid line that carries the inserted gene and then to double the chromosome number either spontaneously or by the use of colchicine. This gives rise to a plant that is homozygous for the inserted gene, which can be easily assayed for if the inserted gene carries with it a suitable selection marker gene for detection of plants carrying that gene. Alternatively, plants may be self-fertilized, leading to the production of a mixture of seed that consists of, in the simplest case, three types, homozygous (25%), heterozygous (50%) and null (25%) for the inserted gene. Although it is relatively easy to score null plants from those that contain the gene, it is possible in practice to score the homozygous from heterozygous plants by southern blot analysis in

which careful attention is paid to the loading of exactly equivalent amounts of DNA from the mixed population, and scoring heterozygotes by the intensity of the signal from a probe specific for the inserted gene. It is advisable to verify the results of the southern blot analysis by allowing each independent transformant to self-fertilize, since additional evidence for homozygosity can be obtained by the simple fact that if the plant was homozygous for the inserted gene, all of the subsequent plants from the selfed seed will contain the gene, while if the plant was heterozygous for the gene, the generation grown from the selfed seed will contain null plants. Therefore, with simple selfing one can easily select homozygous plant lines that can also be confirmed by southern blot analysis.

Creation of homozygous parental lines makes possible the production of hybrid plants and seeds which will contain a modified protein component. Transgenic homozygous parental lines are maintained with each parent containing either the first or second recombinant DNA sequence operably linked to a promoter. Also incorporated in this scheme are the advantages of growing a hybrid crop, including the combining of more valuable traits and hybrid vigor.

In the foregoing discussion, a number of citations from professional journals and patents are included for reference. All such citations are hereby incorporated in their entirety by reference.

EXAMPLE 1

Creation of OPH genes

Utilizing the amino acid sequence disclosed in McDaniel et al., J Bacteriology, 170 (5); 230-11 (May 1988), an OPH gene was created. This was accomplished by analyzing the amino acid sequence of OPH using the program BACKTRANSLATE (Wisconsin package version 9.1, Genetics Computer Group(GCG), Madison, Wisconsin) into a DNA coding sequence using a maize codon preference table tabulated from highly expressed maize genes (http://www.gcg.com/techsupport/data/maize-high.cod). The optimized sequence

was then searched for putative deleterious sequences compiled in Table 1 using the GCG program FINDPATTERNS.

TABLE 1

COMPILATION OF DELETERIOUS mRNA SIGNALS		
NAME	OFFSET	PATTERN
Killer	1	ATTA
mRNA deg	1	AAAA
mRNA deg	1	TTTT
mRNA deg	1	TTTTRTY
AT String	1	W(5,)
Poly A	1	AANNAA
5' Splice	1	VRGTRANN
3' Splice	1	YAGV
mRNA deg	1	ATTTA
mRNA deg	1	AATAA
Poly A	1	AATAAA
Poly A	1	AATAAT
Poly A	1	AACCAA
5' Splice	1	ATAGCAA

Alternative codons were chosen to eliminate the deleterious mRNA signals in the DNA coding sequence and to add convenient restriction enzyme sites to facilitate downstream cloning. Codons that reflected less than 20% usage were avoided. The optimized DNA coding sequence was divided into a convenient cloning length of less than 350 base pairs.

Each of the fragments was divided up into oligonucleotides of approximately 50 bases in length with 10 base overhangs. Oligonucleotides corresponding to both the sense and antisense strand were synthesized using a

commercial supplier and all oligonucleotides were 5' phosphorylated. The oligonucleotides were diluted to 100 pmol/μl and 1 nanomole of each oligonucleotide in each respective fragment was added to a microfuge tube and brought to 100 μl using water. The oligonucleotides were annealed by heating to 95°C and allowed to cool slowly to 4°C over a period of two hours. To 80 μl of the annealed oligonucleotides, 10 μl of 10X ligation buffer (New England Biolabs), and 10 μl of T4 DNA ligase (4,000 units) were added. The oligonucleotides were ligated at 16°C for one hour. Following the ligation, 1 μl and successive 1/10 dilutions were subjected to PCR using primers specific to the 5' and 3' ends of each fragment. The PCR reactions were then purified by polyacrylimide gel electrophoresis. The band corresponding to the correct size of each fragment was excised, and the DNA eluted from the gel slice using proceduresknown to those of skill in the art. The eluted DNA was cloned into the pCR2.1-TOPO (Invitrogen) as per the manufacturer's directions. Colonies containing the PCR insert were grown overnight, the plasmid DNA extracted and purified, and the plasmid DNA analyzed for the correct PCR insert as per standard protocols to those known of skill in the art. The inserts were sequenced to verify the correct sequence of each fragment and then the optimized gene was assembled (SEQ ID NO:1) using standard protocols known to those of skill in the art. The 5' fragment was assembled to contain a plant leader sequence (SEQ ID NO:2) to ensure high expression in plants and the OPH gene containing the plant leader (SEQ ID NO:3) was cloned into a standard plant expression cassette to ensure high expression in plants.

EXAMPLE 2

Plant Transformation

While it is to be understood that any method of plant transformation known to those of skill in the art may be used on a large variety of plants, the following example describes plant transformation by electroporation. The procedure used to electroporate plant protoplasts is essentially as described by David Chen and coworkers in the Hoefer Scientific Instruments Technical

Bulletin No. 118. The upper epidermis of plant leaves is isolated when the plant leaf is 3 to 4 cm in length and then leaf tissue is brushed with 320 grit aluminum oxide powder to permit the infiltration of cell wall degradative enzymes used to prepare protoplasts by the method of Magnien, E. et al., Acta Genetica Sinica 7, 231 (1980). Enzymatically released protoplasts are washed with 17.5% sucrose, floated and harvested by centrifugation for 5 minutes at 300 x g in a 60 ml Babcock bottle. Linearized or supercoiled plant vector DNA containing the OPH gene and plant leader sequence is mixed with the protoplasts in a final volume of 0.5 ml at a concentration of 0.1 mg/ml and 7 x 10^5 cells/ml respectively in a 16 mm diameter Nunc Multidish well. A single pulse is administered at room temperature (23°C) in a Hoefer PG 101 ProGenetor electroporation unit using a PG120-2.5 electrode for 10 msec at 200 V. Electroporated protoplasts are kept stationary for 10 minutes prior to the addition of 1 ml of culture medium. Cells were subsequently diluted to a final concentration of 10^5 cells/ml. These cells may then be assayed for transient expression of the OPH gene after a period of 40-48 hours or depending on the DNA vector construct used, plated to generate callus tissues under appropriate selection, followed by regeneration to whole plants.

EXAMPLE 3

Regeneration of Maize Plants from Somatic Cell Cultures

While it is to be understood that many different types and varieties of plants may be transformed and ultimately regenerated by methods known to those of skill in the art, this example describes maize plant regeneration from somatic cell cultures, specifically Type II cultures. The procedure used to regenerate maize plant from Type II cultures is essentially as described in The Maize Handbook (Freeling & Walbot, eds.) (Springer-Verlag pub. 1994) at pages 667-70. Regeneration of Type II cultures is accomplished in a three-step process. This three-step process utilizes three differing media: Regen I, Regen II, and Regen III (the media recipes can be found at page 668 of Freeling & Walbot, The Maize Handbook). Regen I promotes somatic embryo

differentiation (See Abbe & Stein, "The Origin of the Shoot Apex in Maize; Embryogeny", N.J. Bot. 41:285, 293 (1954)). Regen II promotes embryo enlargement and maturation, and Regen III promotes germination. The tissue is incubated on each medium for approximately two weeks. The first two incubation steps are carried out in the dark at 28°C and the final incubation step is carried out under a 16:8-hour photo period with about 70 $\mu E\ m^{-2}\ sec^{-1}$ provided by cool-white fluorescent bulbs at about 25°C. Upon completion of all three incubations, the plates are placed in a single layer on shelves in a lighted incubator and are sealed with 3-M Scotch Brand #394 mending tape. 3-M tape is preferred over either Parafilm or masking tape as it "breathes" better and helps reduce condensation problems. Small green shoots formed on Regen III in 100 x 25-mm petri plates are transferred to Regen III medium and 200 x 25-mm Pyrex tubes or Phytatrays to permit further plantlet development and root formation. Once a good root system has developed, the plants may be carefully removed from the medium, the root system washed thoroughly under running water, and the plants placed into a 2. 5" square pot containing Metromix 350 growing medium. Watering with the systemic fungicide benomyl (100 mg/1 Benlate) will reduce the risk of fungal contamination. It is critical to maintain the freshly transplanted plants in a high humidity environment for several days. Placing flats full of well-watered plants into large, sealed, transparent plastic bags is a simple, economical way to accomplish this. Condensation should be visible on the inside of the plastic bags. After several days, the humidity should be gradually reduced to harden off the plants. This can be done by cutting a few holes in the bag and watering the plants again at this point if needed. Several days later (about one week since taking to soil), the plant should be hardened off well enough to survive typical growth chamber or greenhouse conditions for maize. Transplanting from 2.5-inch square to 6-inch diameter round and finally to 10-12-inch diameter round pots works well. Metromix 350 or a similar growing medium works well for all stages, however, the plants still should be fertilized. In addition, during the hardening off stage only, temperature and light intensity should be lower than normal for maize

greenhouse conditions (for example, 23°C and 2-5 μE m^{-2} set^{-1} works well for hardening off).

EXAMPLE 4

Maize Plant Regeneration From Type II Callus

Regeneration of plants from Type II callus is based upon allowing the embryoids on the surface of the Type II callus to mature and germinate. See Freeling & Walbot, The Maize Handbook (1994) at page 673-674. The procedure used is essentially that as described in Freeling & Walbot. The callus are first collected and weighed in petri plates. 1-2 grams fresh weight of soft, friable Type II callus containing numerous embryoids are evenly distributed over the surface of a 100 x 15-mm petri plate which contains 25 ml of regeneration medium. Regeneration medium consists of Murashige and Skoog (MS) basal salts, modified White's vitamins (0.2 g/l glycine, and 0.5 g/l of each of thiamine-HCl, pyridoxine-HCl, and nicotinic acid), supplemented with 6% sucrose, 0.1 g/l myo-inositol, and 0.8% Bacto-ager (6SMS0D). The plates are then wrapped with Parafilm and placed in the dark. After one week, the plates are moved to a lighted growth chamber with a 16 hour (75 μE m^{-2} sec^{-1}) and an 8 hour dark photo period. Three weeks after plating the Type II callus to the 6SMS0D, the plate should be examined for shoot formation from the calli. The calli and the shoot should then be transferred to fresh Bacto-ager plates for another two weeks. The callus without shoots can be left on the Bacto-ager for a longer period if the callus is slow in embryo development. Upon distinct formation of a shoot and root (some may be ready for transfer after the first three weeks on regeneration medium) transfer the newly developed green plantlets to Magenta GA-7 (Magenta Corp., Chicago, IL) containers containing 60 ml of 3SMSOD medium solidified with 0.6% Bacto-ager. When the plant has developed a strong root system (10-15 days after transfer into Magenta boxes) gently remove the plant from the ager, wash the remaining ager from the roots and shoot and carefully transplant the plant into a 4-inch pot containing a moist soil. The pots can then be placed in

a high humidity chamber and, over a period of ten days, slowly reduce the humidity to approximately that of the greenhouse. Once plants are adapted to a lower humidity, they may be removed a greenhouse and treated like seedlings.

EXAMPLE 5

Production and Stability of Optimized OPH in Transgenic Maize

Organophosphorous hydrolase that is produced by transformed maize plants containing the optimized OPH gene and plant leader sequence is purified by standard methods known to those in the art. The presence of the OPH protein can be detected utilizing protein purified from transformed plant tissues. Further, the expression and stability of the OPH gene product can be assayed by allowing the OPH optimized gene product to come into association with the pesticide, neurotoxin, or compound containing phosphoryl-bonds. These compounds can then be analyzed to determine if the OPH gene product hydrolyzed the P-O, P-F, P-CN, and/or P-S phosphoryl-bonds.

One of ordinary skill in the art, with the aid of the present disclosure, can affect various changes, substitutions of equivalents and other alterations to the methods and compositions herein set forth, in order to practice this invention. Therefore, the protection granted by letters patent should not be limited except by the language of the claims as set forth below.

What is claimed is:

1. A nucleotide sequence which encodes upon expression an organophosphorous hydrolase protein, said nucleotide sequence optimized for plant expression.

2. The nucleotide sequence of claim 1 wherein said sequence is capable of achieving expression levels greater than 10-15 mg/l, typically in bacterial fermentation.

3. The nucleotide sequence of claim 1 wherein said sequence is capable of hybridizing under conditions of high stringency to SEQ ID NO:3.

4. The nucleotide sequence of claim 1 wherein said sequence is SEQ ID NO:3.

5. The nucleotide sequence of claim 1 wherein said sequence is capable of hybridizing under conditions of high stringency to SEQ ID NO:5.

6. The nucleotide sequence of claim 1 wherein said sequence is SEQ ID NO:5.

7. An expression construct comprising: a nucleotide sequence according to claim 1 operably linked to a regulatory region capable of directing expression of said sequence in a plant cell.

8. The expression construct of claim 7 wherein said regulatory region includes a leader sequence.

9. The expression construct of claim 8 wherein said leader sequence is SEQ ID NO:1.

10. The expression construct of claim 7 wherein said nucleotide sequence is capable of hybridyzing under conditions of high strigency to SEQ ID NO:3.

11. The expression construct of claim 7 wherein said nucleotide sequence is SEQ ID NO:1.

12. The expression construction of claim 7 wherein said regulatory region comprises a promoter.

13. The expression construct of claim 7 wherein said promoter is selected from the group consisting of:
a constitutive promoter, an inducible promoter, and a tissue specific promoter.

14. The construct of claim 12 wherein said promoter is a seed specific promoter.

15. A vector capable of transforming or transfecting a host plant cell, said vector comprising an expression construct according to claim 7.

16. The vector of claim 15 wherein said vector is a plasmid based vector.

17. The vector of claim 15 wherein said vector is a viral based vector.

18. A procaryotic or eukaryotic host cell transformed or transfected with the vector of claim 15.

19. The host cell of claim 18 wherein said cell is a plant cell.

20. An organophosphorous hydrolase protein produced by the method of:
obtaining a plant optimized organophosphorous hydrolase protein encoding

nucleic acid sequence; transforming a plant cell with said sequence to create a transgenic plant which is capable of expressing said optimized sequence; and thereafter harvesting said protein from said transgenic plant.

21. The protein of claim 20 wherein said organophosphorous hydrolase protein encoding sequence is one which is capable of hybridizing under conditions of high strigency to SEQ ID NO:3.

22. The protein of claim 21 where said organophosphorous hydrolase protein encoding sequence is SEQ ID NO:3.

23. A plant which is capable of expressing organophosphorous hydrolase protein.

24. A method of environmental and *in situ* detoxification comprising: exposing said contaminated material to organophosphorous hydrolase protein produced by the method of claim 16.

25. A method of environmental detoxification of organophosphorous neurotoxin contaminated soil comprising: introducing to said soil a transgenic plant according to claim 19.

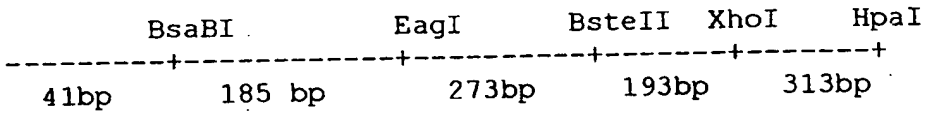


Figure 1

FIGURE 2

```
cc atg gcc aac aag cac ctg agc ctc tcc ctc ttc ctc gtg ctc ctc     47
   Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu
     1               5                  10                  15

ggc ctc tcc gcc tcc ctc gcc agc                                   71
Gly Leu Ser Ala Ser Leu Ala Ser
                20
```

FIGURE 3A

```
ggc acc ggc gac cgc atc aac acc gtg cgc ggc ccg atc acc atc tcc   48
Gly Thr Gly Asp Arg Ile Asn Thr Val Arg Gly Pro Ile Thr Ile Ser
  1               5                  10                  15

gag gcc ggc ttc acc ctc acc cac gag cac atc tgc ggc tcc tcc gcc   96
Glu Ala Gly Phe Thr Leu Thr His Glu His Ile Cys Gly Ser Ser Ala
             20                  25                  30

ggc ttc ctc cgc gcc tgg ccg gag ttc ttc ggc tcc cgc aag gcc ctc   144
Gly Phe Leu Arg Ala Trp Pro Glu Phe Phe Gly Ser Arg Lys Ala Leu
         35                  40                  45

gcc gag aag gcc gtg cgc ggc ctc cgc cgc gcc cgc gcc gcc ggc gtg   192
Ala Glu Lys Ala Val Arg Gly Leu Arg Arg Ala Arg Ala Ala Gly Val
     50                  55                  60

cgc acc atc gtg gac gtg tcc acc ttc gac atc ggc cgc gac gtg tcc   240
Arg Thr Ile Val Asp Val Ser Thr Phe Asp Ile Gly Arg Asp Val Ser
 65                  70                  75                  80

ctc ctc gcc gag gtg tcc cgc gcc gcc gac gtg cac atc gtg gcc gcc   288
Leu Leu Ala Glu Val Ser Arg Ala Ala Asp Val His Ile Val Ala Ala
                 85                  90                  95

acc ggc ctc tgg ttc gac ccg ccg ctc tcc atg cgc ctc cgc tcc gtg   336
Thr Gly Leu Trp Phe Asp Pro Pro Leu Ser Met Arg Leu Arg Ser Val
            100                 105                 110

gag gag ctc acc cag ttc ttc ctc cgc gag atc cag tac ggc atc gag   384
Glu Glu Leu Thr Gln Phe Phe Leu Arg Glu Ile Gln Tyr Gly Ile Glu
        115                 120                 125

gac acc ggc atc cgc gcc ggc atc atc aag gtg gcc acc acc ggc aag   432
Asp Thr Gly Ile Arg Ala Gly Ile Ile Lys Val Ala Thr Thr Gly Lys
    130                 135                 140

gcc acc ccg ttc cag gag ctc gtg ctc aag gcc gcc gcc cgc gcc tcc   480
Ala Thr Pro Phe Gln Glu Leu Val Leu Lys Ala Ala Ala Arg Ala Ser
145                 150                 155                 160

ctc gcc acc ggc gtg ccg gtg acc acc cac acc gcc gcc tcc cag cgc   528
Leu Ala Thr Gly Val Pro Val Thr Thr His Thr Ala Ala Ser Gln Arg
                165                 170                 175

gac ggc gag cag cag gcc gcc atc ttc gag tcc gag ggc ctc tcc ccg   576
Asp Gly Glu Gln Gln Ala Ala Ile Phe Glu Ser Glu Gly Leu Ser Pro
            180                 185                 190

tcc cgc gtg tgc atc ggc cac tcc gac gac acc gac gac ctc tcc tac   624
Ser Arg Val Cys Ile Gly His Ser Asp Asp Thr Asp Asp Leu Ser Tyr
        195                 200                 205

ctc acc gcc ctc gcc gcc cgc ggc tac ctc atc ggc ctc gac cac atc   672
Leu Thr Ala Leu Ala Ala Arg Gly Tyr Leu Ile Gly Leu Asp His Ile
    210                 215                 220
```

FIGURE 3B

```
ccg cac tcc gcc atc ggc ctc gag gac aac gcc tcc gcc tcc gcc ctc   720
Pro His Ser Ala Ile Gly Leu Glu Asp Asn Ala Ser Ala Ser Ala Leu
225                 230                 235                 240

ctc ggc atc cgc tcc tgg cag acc cgc gcc ctc ctc atc aag gcc ctc   768
Leu Gly Ile Arg Ser Trp Gln Thr Arg Ala Leu Leu Ile Lys Ala Leu
                245                 250                 255

atc gac cag ggc tac atg aag cag atc ctc gtg tcc aac gac tgg ctc   816
Ile Asp Gln Gly Tyr Met Lys Gln Ile Leu Val Ser Asn Asp Trp Leu
            260                 265                 270

ttc ggc ttc tcc tcc tac gtg acc aac atc atg gac gtg atg gac cgc   864
Phe Gly Phe Ser Ser Tyr Val Thr Asn Ile Met Asp Val Met Asp Arg
        275                 280                 285

gtg aac ccg gac ggc atg gcc ttc atc ccg ctc cgc gtg atc ccg ttc   912
Val Asn Pro Asp Gly Met Ala Phe Ile Pro Leu Arg Val Ile Pro Phe
    290                 295                 300

ctc cgc gag aag ggc gtg ccg cag gag acc ctc gcc ggc atc acc gtg   960
Leu Arg Glu Lys Gly Val Pro Gln Glu Thr Leu Ala Gly Ile Thr Val
305                 310                 315                 320

acc aac ccg gcc cgc ttc ctc tcc ccg acc ctc cgc gcc tcc tga      1005
Thr Asn Pro Ala Arg Phe Leu Ser Pro Thr Leu Arg Ala Ser
                325                 330                 335
```

FIGURE 4A

```
cc atg gcc aac aag cac ctg agc ctc tcc ctc ttc ctc gtg ctc ctc    47
   Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu
     1               5                  10                  15

ggc ctc tcc gcc tcc ctc gcc agc ggc acc ggc gac cgc atc aac acc   95
Gly Leu Ser Ala Ser Leu Ala Ser Gly Thr Gly Asp Arg Ile Asn Thr
                 20                  25                  30

gtg cgc ggc ccg atc acc atc tcc gag gcc ggc ttc acc ctc acc cac   143
Val Arg Gly Pro Ile Thr Ile Ser Glu Ala Gly Phe Thr Leu Thr His
             35                  40                  45

gag cac atc tgc ggc tcc tcc gcc ggc ttc ctc cgc gcc tgg ccg gag   191
Glu His Ile Cys Gly Ser Ser Ala Gly Phe Leu Arg Ala Trp Pro Glu
         50                  55                  60

ttc ttc ggc tcc cgc aag gcc ctc gcc gag aag gcc gtg cgc ggc ctc   239
Phe Phe Gly Ser Arg Lys Ala Leu Ala Glu Lys Ala Val Arg Gly Leu
     65                  70                  75

cgc cgc gcc cgc gcc gcc ggc gtg cgc acc atc gtg gac gtg tcc acc   287
Arg Arg Ala Arg Ala Ala Gly Val Arg Thr Ile Val Asp Val Ser Thr
 80                  85                  90                  95

ttc gac atc ggc cgc gac gtg tcc ctc ctc gcc gag gtg tcc cgc gcc   335
Phe Asp Ile Gly Arg Asp Val Ser Leu Leu Ala Glu Val Ser Arg Ala
                100                 105                 110

gcc gac gtg cac atc gtg gcc gcc acc ggc ctc tgg ttc gac ccg ccg   383
Ala Asp Val His Ile Val Ala Ala Thr Gly Leu Trp Phe Asp Pro Pro
            115                 120                 125

ctc tcc atg cgc ctc cgc tcc gtg gag gag ctc acc cag ttc ttc ctc   431
Leu Ser Met Arg Leu Arg Ser Val Glu Glu Leu Thr Gln Phe Phe Leu
        130                 135                 140

cgc gag atc cag tac ggc atc gag gac acc ggc atc cgc gcc ggc atc   479
Arg Glu Ile Gln Tyr Gly Ile Glu Asp Thr Gly Ile Arg Ala Gly Ile
    145                 150                 155

atc aag gtg gcc acc acc ggc aag gcc acc ccg ttc cag gag ctc gtg   527
Ile Lys Val Ala Thr Thr Gly Lys Ala Thr Pro Phe Gln Glu Leu Val
160                 165                 170                 175

ctc aag gcc gcc gcc cgc gcc tcc ctc gcc acc ggc gtg ccg gtg acc   575
Leu Lys Ala Ala Ala Arg Ala Ser Leu Ala Thr Gly Val Pro Val Thr
                180                 185                 190

acc cac acc gcc gcc tcc cag cgc gac ggc gag cag cag gcc gcc atc   623
Thr His Thr Ala Ala Ser Gln Arg Asp Gly Glu Gln Gln Ala Ala Ile
            195                 200                 205

ttc gag tcc gag ggc ctc tcc ccg tcc cgc gtg tgc atc ggc cac tcc   671
Phe Glu Ser Glu Gly Leu Ser Pro Ser Arg Val Cys Ile Gly His Ser
        210                 215                 220
```

FIGURE 4B

```
gac gac acc gac gac ctc tcc tac ctc acc gcc ctc gcc gcc cgc ggc   719
Asp Asp Thr Asp Asp Leu Ser Tyr Leu Thr Ala Leu Ala Ala Arg Gly
    225                 230                 235

tac ctc atc ggc ctc gac cac atc ccg cac tcc gcc atc ggc ctc gag   767
Tyr Leu Ile Gly Leu Asp His Ile Pro His Ser Ala Ile Gly Leu Glu
240                 245                 250                 255

gac aac gcc tcc gcc tcc gcc ctc ctc ggc atc cgc tcc tgg cag acc   815
Asp Asn Ala Ser Ala Ser Ala Leu Leu Gly Ile Arg Ser Trp Gln Thr
                260                 265                 270

cgc gcc ctc ctc atc aag gcc ctc atc gac cag ggc tac atg aag cag   863
Arg Ala Leu Leu Ile Lys Ala Leu Ile Asp Gln Gly Tyr Met Lys Gln
            275                 280                 285

atc ctc gtg tcc aac gac tgg ctc ttc ggc ttc tcc tcc tac gtg acc   911
Ile Leu Val Ser Asn Asp Trp Leu Phe Gly Phe Ser Ser Tyr Val Thr
        290                 295                 300

aac atc atg gac gtg atg gac cgc gtg aac ccg gac ggc atg gcc ttc   959
Asn Ile Met Asp Val Met Asp Arg Val Asn Pro Asp Gly Met Ala Phe
    305                 310                 315

atc ccg ctc cgc gtg atc ccg ttc ctc cgc gag aag ggc gtg ccg cag   1007
Ile Pro Leu Arg Val Ile Pro Phe Leu Arg Glu Lys Gly Val Pro Gln
320                 325                 330                 335

gag acc ctc gcc ggc atc acc gtg acc aac ccg gcc cgc ttc ctc tcc   1055
Glu Thr Leu Ala Gly Ile Thr Val Thr Asn Pro Ala Arg Phe Leu Ser
                340                 345                 350

ccg acc ctc cgc gcc tcc tga                                     1076
Pro Thr Leu Arg Ala Ser
            355
```

SEQUENCE LISTING

<110> Jilka, Joseph M.

<120> Optimized Nucleotide Sequence Encoding
Organophosphorous Hydrolase and Methods of Use for Same

<130> jilka

<140>
<141>

<150> 60/081,846
<151> 1998-04-15

<160> 6

<170> PatentIn Ver. 2.0

<210> 1
<211> 71
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: barley alpha
amylase leader sequence

<220>
<221> CDS
<222> (3)..(71)

<400> 1
cc atg gcc aac aag cac ctg agc ctc tcc ctc ttc ctc gtg ctc ctc 47
Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu
1 5 10 15

ggc ctc tcc gcc tcc ctc gcc agc 71
Gly Leu Ser Ala Ser Leu Ala Ser
20

<210> 2
<211> 23
<212> PRT
<213> Artificial Sequence

<400> 2
Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

Leu Ser Ala Ser Leu Ala Ser
20

<210> 3
<211> 1005
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: OPH optimized

<220>
<221> CDS
<222> (1)..(1005)

<400> 3

```
ggc acc ggc gac cgc atc aac acc gtg cgc ggc ccg atc acc atc tcc   48
Gly Thr Gly Asp Arg Ile Asn Thr Val Arg Gly Pro Ile Thr Ile Ser
  1               5                  10                  15

gag gcc ggc ttc acc ctc acc cac gag cac atc tgc ggc tcc tcc gcc   96
Glu Ala Gly Phe Thr Leu Thr His Glu His Ile Cys Gly Ser Ser Ala
             20                  25                  30

ggc ttc ctc cgc gcc tgg ccg gag ttc ttc ggc tcc cgc aag gcc ctc   144
Gly Phe Leu Arg Ala Trp Pro Glu Phe Phe Gly Ser Arg Lys Ala Leu
         35                  40                  45

gcc gag aag gcc gtg cgc ggc ctc cgc cgc gcc cgc gcc gcc ggc gtg   192
Ala Glu Lys Ala Val Arg Gly Leu Arg Arg Ala Arg Ala Ala Gly Val
     50                  55                  60

cgc acc atc gtg gac gtg tcc acc ttc gac atc ggc cgc gac gtg tcc   240
Arg Thr Ile Val Asp Val Ser Thr Phe Asp Ile Gly Arg Asp Val Ser
 65                  70                  75                  80

ctc ctc gcc gag gtg tcc cgc gcc gcc gac gtg cac atc gtg gcc gcc   288
Leu Leu Ala Glu Val Ser Arg Ala Ala Asp Val His Ile Val Ala Ala
                 85                  90                  95

acc ggc ctc tgg ttc gac ccg ccg ctc tcc atg cgc ctc cgc tcc gtg   336
Thr Gly Leu Trp Phe Asp Pro Pro Leu Ser Met Arg Leu Arg Ser Val
            100                 105                 110

gag gag ctc acc cag ttc ttc ctc cgc gag atc cag tac ggc atc gag   384
Glu Glu Leu Thr Gln Phe Phe Leu Arg Glu Ile Gln Tyr Gly Ile Glu
        115                 120                 125

gac acc ggc atc cgc gcc ggc atc atc aag gtg gcc acc acc ggc aag   432
Asp Thr Gly Ile Arg Ala Gly Ile Ile Lys Val Ala Thr Thr Gly Lys
    130                 135                 140

gcc acc ccg ttc cag gag ctc gtg ctc aag gcc gcc gcc cgc gcc tcc   480
Ala Thr Pro Phe Gln Glu Leu Val Leu Lys Ala Ala Ala Arg Ala Ser
145                 150                 155                 160

ctc gcc acc ggc gtg ccg gtg acc acc cac acc gcc gcc tcc cag cgc   528
Leu Ala Thr Gly Val Pro Val Thr Thr His Thr Ala Ala Ser Gln Arg
                165                 170                 175

gac ggc gag cag cag gcc gcc atc ttc gag tcc gag ggc ctc tcc ccg   576
Asp Gly Glu Gln Gln Ala Ala Ile Phe Glu Ser Glu Gly Leu Ser Pro
            180                 185                 190

tcc cgc gtg tgc atc ggc cac tcc gac gac acc gac gac ctc tcc tac   624
Ser Arg Val Cys Ile Gly His Ser Asp Asp Thr Asp Asp Leu Ser Tyr
```

```
        195                 200                 205

ctc acc gcc ctc gcc gcc cgc ggc tac ctc atc ggc ctc gac cac atc   672
Leu Thr Ala Leu Ala Ala Arg Gly Tyr Leu Ile Gly Leu Asp His Ile
    210                 215                 220

ccg cac tcc gcc atc ggc ctc gag gac aac gcc tcc gcc tcc gcc ctc   720
Pro His Ser Ala Ile Gly Leu Glu Asp Asn Ala Ser Ala Ser Ala Leu
225                 230                 235                 240

ctc ggc atc cgc tcc tgg cag acc cgc gcc ctc ctc atc aag gcc ctc   768
Leu Gly Ile Arg Ser Trp Gln Thr Arg Ala Leu Leu Ile Lys Ala Leu
                245                 250                 255

atc gac cag ggc tac atg aag cag atc ctc gtg tcc aac gac tgg ctc   816
Ile Asp Gln Gly Tyr Met Lys Gln Ile Leu Val Ser Asn Asp Trp Leu
            260                 265                 270

ttc ggc ttc tcc tcc tac gtg acc aac atc atg gac gtg atg gac cgc   864
Phe Gly Phe Ser Ser Tyr Val Thr Asn Ile Met Asp Val Met Asp Arg
        275                 280                 285

gtg aac ccg gac ggc atg gcc ttc atc ccg ctc cgc gtg atc ccg ttc   912
Val Asn Pro Asp Gly Met Ala Phe Ile Pro Leu Arg Val Ile Pro Phe
    290                 295                 300

ctc cgc gag aag ggc gtg ccg cag gag acc ctc gcc ggc atc acc gtg   960
Leu Arg Glu Lys Gly Val Pro Gln Glu Thr Leu Ala Gly Ile Thr Val
305                 310                 315                 320

acc aac ccg gcc cgc ttc ctc tcc ccg acc ctc cgc gcc tcc tga      1005
Thr Asn Pro Ala Arg Phe Leu Ser Pro Thr Leu Arg Ala Ser
                325                 330                 335
```

<210> 4
<211> 334
<212> PRT
<213> Artificial Sequence

<400> 4

```
Gly Thr Gly Asp Arg Ile Asn Thr Val Arg Gly Pro Ile Thr Ile Ser
  1               5                  10                  15

Glu Ala Gly Phe Thr Leu Thr His Glu His Ile Cys Gly Ser Ser Ala
             20                  25                  30

Gly Phe Leu Arg Ala Trp Pro Glu Phe Phe Gly Ser Arg Lys Ala Leu
         35                  40                  45

Ala Glu Lys Ala Val Arg Gly Leu Arg Arg Ala Arg Ala Ala Gly Val
     50                  55                  60

Arg Thr Ile Val Asp Val Ser Thr Phe Asp Ile Gly Arg Asp Val Ser
 65                  70                  75                  80

Leu Leu Ala Glu Val Ser Arg Ala Ala Asp Val His Ile Val Ala Ala
                 85                  90                  95
```

```
Thr Gly Leu Trp Phe Asp Pro Pro Leu Ser Met Arg Leu Arg Ser Val
            100                 105                 110

Glu Glu Leu Thr Gln Phe Phe Leu Arg Glu Ile Gln Tyr Gly Ile Glu
        115                 120                 125

Asp Thr Gly Ile Arg Ala Gly Ile Ile Lys Val Ala Thr Thr Gly Lys
    130                 135                 140

Ala Thr Pro Phe Gln Glu Leu Val Leu Lys Ala Ala Ala Arg Ala Ser
145                 150                 155                 160

Leu Ala Thr Gly Val Pro Val Thr Thr His Thr Ala Ala Ser Gln Arg
                165                 170                 175

Asp Gly Glu Gln Gln Ala Ala Ile Phe Glu Ser Glu Gly Leu Ser Pro
            180                 185                 190

Ser Arg Val Cys Ile Gly His Ser Asp Asp Thr Asp Asp Leu Ser Tyr
        195                 200                 205

Leu Thr Ala Leu Ala Ala Arg Gly Tyr Leu Ile Gly Leu Asp His Ile
    210                 215                 220

Pro His Ser Ala Ile Gly Leu Glu Asp Asn Ala Ser Ala Ser Ala Leu
225                 230                 235                 240

Leu Gly Ile Arg Ser Trp Gln Thr Arg Ala Leu Leu Ile Lys Ala Leu
                245                 250                 255

Ile Asp Gln Gly Tyr Met Lys Gln Ile Leu Val Ser Asn Asp Trp Leu
            260                 265                 270

Phe Gly Phe Ser Ser Tyr Val Thr Asn Ile Met Asp Val Met Asp Arg
        275                 280                 285

Val Asn Pro Asp Gly Met Ala Phe Ile Pro Leu Arg Val Ile Pro Phe
    290                 295                 300

Leu Arg Glu Lys Gly Val Pro Gln Glu Thr Leu Ala Gly Ile Thr Val
305                 310                 315                 320

Thr Asn Pro Ala Arg Phe Leu Ser Pro Thr Leu Arg Ala Ser
                325                 330
```

<210> 5
<211> 1076
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: optimized OPH plus leader

<220>
<221> CDS
<222> (3)..(1076)

<400> 5

```
cc atg gcc aac aag cac ctg agc ctc tcc ctc ttc ctc gtg ctc ctc    47
   Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu
     1               5                  10                  15

ggc ctc tcc gcc tcc ctc gcc agc ggc acc ggc gac cgc atc aac acc   95
Gly Leu Ser Ala Ser Leu Ala Ser Gly Thr Gly Asp Arg Ile Asn Thr
                 20                  25                  30

gtg cgc ggc ccg atc acc atc tcc gag gcc ggc ttc acc ctc acc cac   143
Val Arg Gly Pro Ile Thr Ile Ser Glu Ala Gly Phe Thr Leu Thr His
             35                  40                  45

gag cac atc tgc ggc tcc tcc gcc ggc ttc ctc cgc gcc tgg ccg gag   191
Glu His Ile Cys Gly Ser Ser Ala Gly Phe Leu Arg Ala Trp Pro Glu
         50                  55                  60

ttc ttc ggc tcc cgc aag gcc ctc gcc gag aag gcc gtg cgc ggc ctc   239
Phe Phe Gly Ser Arg Lys Ala Leu Ala Glu Lys Ala Val Arg Gly Leu
     65                  70                  75

cgc cgc gcc cgc gcc gcc ggc gtg cgc acc atc gtg gac gtg tcc acc   287
Arg Arg Ala Arg Ala Ala Gly Val Arg Thr Ile Val Asp Val Ser Thr
 80                  85                  90                  95

ttc gac atc ggc cgc gac gtg tcc ctc ctc gcc gag gtg tcc cgc gcc   335
Phe Asp Ile Gly Arg Asp Val Ser Leu Leu Ala Glu Val Ser Arg Ala
                100                 105                 110

gcc gac gtg cac atc gtg gcc gcc acc ggc ctc tgg ttc gac ccg ccg   383
Ala Asp Val His Ile Val Ala Ala Thr Gly Leu Trp Phe Asp Pro Pro
            115                 120                 125

ctc tcc atg cgc ctc cgc tcc gtg gag gag ctc acc cag ttc ttc ctc   431
Leu Ser Met Arg Leu Arg Ser Val Glu Glu Leu Thr Gln Phe Phe Leu
        130                 135                 140

cgc gag atc cag tac ggc atc gag gac acc ggc atc cgc gcc ggc atc   479
Arg Glu Ile Gln Tyr Gly Ile Glu Asp Thr Gly Ile Arg Ala Gly Ile
    145                 150                 155

atc aag gtg gcc acc acc ggc aag gcc acc ccg ttc cag gag ctc gtg   527
Ile Lys Val Ala Thr Thr Gly Lys Ala Thr Pro Phe Gln Glu Leu Val
160                 165                 170                 175

ctc aag gcc gcc gcc cgc gcc tcc ctc gcc acc ggc gtg ccg gtg acc   575
Leu Lys Ala Ala Ala Arg Ala Ser Leu Ala Thr Gly Val Pro Val Thr
                180                 185                 190

acc cac acc gcc gcc tcc cag cgc gac ggc gag cag cag gcc gcc atc   623
Thr His Thr Ala Ala Ser Gln Arg Asp Gly Glu Gln Gln Ala Ala Ile
            195                 200                 205

ttc gag tcc gag ggc ctc tcc ccg tcc cgc gtg tgc atc ggc cac tcc   671
Phe Glu Ser Glu Gly Leu Ser Pro Ser Arg Val Cys Ile Gly His Ser
        210                 215                 220

gac gac acc gac gac ctc tcc tac ctc acc gcc ctc gcc gcc cgc ggc   719
Asp Asp Thr Asp Asp Leu Ser Tyr Leu Thr Ala Leu Ala Ala Arg Gly
```

```
    225                 230                 235

tac ctc atc ggc ctc gac cac atc ccg cac tcc gcc atc ggc ctc gag   767
Tyr Leu Ile Gly Leu Asp His Ile Pro His Ser Ala Ile Gly Leu Glu
240                 245                 250                 255

gac aac gcc tcc gcc tcc gcc ctc ctc ggc atc cgc tcc tgg cag acc   815
Asp Asn Ala Ser Ala Ser Ala Leu Leu Gly Ile Arg Ser Trp Gln Thr
                260                 265                 270

cgc gcc ctc ctc atc aag gcc ctc atc gac cag ggc tac atg aag cag   863
Arg Ala Leu Leu Ile Lys Ala Leu Ile Asp Gln Gly Tyr Met Lys Gln
            275                 280                 285

atc ctc gtg tcc aac gac tgg ctc ttc ggc ttc tcc tcc tac gtg acc   911
Ile Leu Val Ser Asn Asp Trp Leu Phe Gly Phe Ser Ser Tyr Val Thr
        290                 295                 300

aac atc atg gac gtg atg gac cgc gtg aac ccg gac ggc atg gcc ttc   959
Asn Ile Met Asp Val Met Asp Arg Val Asn Pro Asp Gly Met Ala Phe
    305                 310                 315

atc ccg ctc cgc gtg atc ccg ttc ctc cgc gag aag ggc gtg ccg cag   1007
Ile Pro Leu Arg Val Ile Pro Phe Leu Arg Glu Lys Gly Val Pro Gln
320                 325                 330                 335

gag acc ctc gcc ggc atc acc gtg acc aac ccg gcc cgc ttc ctc tcc   1055
Glu Thr Leu Ala Gly Ile Thr Val Thr Asn Pro Ala Arg Phe Leu Ser
                340                 345                 350

ccg acc ctc cgc gcc tcc tga                                      1076
Pro Thr Leu Arg Ala Ser
            355
```

<210> 6
<211> 357
<212> PRT
<213> Artificial Sequence

<400> 6

```
Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
  1               5                  10                  15

Leu Ser Ala Ser Leu Ala Ser Gly Thr Gly Asp Arg Ile Asn Thr Val
             20                  25                  30

Arg Gly Pro Ile Thr Ile Ser Glu Ala Gly Phe Thr Leu Thr His Glu
         35                  40                  45

His Ile Cys Gly Ser Ser Ala Gly Phe Leu Arg Ala Trp Pro Glu Phe
     50                  55                  60

Phe Gly Ser Arg Lys Ala Leu Ala Glu Lys Ala Val Arg Gly Leu Arg
 65                  70                  75                  80

Arg Ala Arg Ala Ala Gly Val Arg Thr Ile Val Asp Val Ser Thr Phe
                 85                  90                  95
```

```
Asp Ile Gly Arg Asp Val Ser Leu Leu Ala Glu Val Ser Arg Ala Ala
            100                 105                 110

Asp Val His Ile Val Ala Ala Thr Gly Leu Trp Phe Asp Pro Pro Leu
        115                 120                 125

Ser Met Arg Leu Arg Ser Val Glu Glu Leu Thr Gln Phe Phe Leu Arg
    130                 135                 140

Glu Ile Gln Tyr Gly Ile Glu Asp Thr Gly Ile Arg Ala Gly Ile Ile
145                 150                 155                 160

Lys Val Ala Thr Thr Gly Lys Ala Thr Pro Phe Gln Glu Leu Val Leu
                165                 170                 175

Lys Ala Ala Ala Arg Ala Ser Leu Ala Thr Gly Val Pro Val Thr Thr
            180                 185                 190

His Thr Ala Ala Ser Gln Arg Asp Gly Glu Gln Gln Ala Ala Ile Phe
        195                 200                 205

Glu Ser Glu Gly Leu Ser Pro Ser Arg Val Cys Ile Gly His Ser Asp
    210                 215                 220

Asp Thr Asp Asp Leu Ser Tyr Leu Thr Ala Leu Ala Ala Arg Gly Tyr
225                 230                 235                 240

Leu Ile Gly Leu Asp His Ile Pro His Ser Ala Ile Gly Leu Glu Asp
                245                 250                 255

Asn Ala Ser Ala Ser Ala Leu Leu Gly Ile Arg Ser Trp Gln Thr Arg
            260                 265                 270

Ala Leu Leu Ile Lys Ala Leu Ile Asp Gln Gly Tyr Met Lys Gln Ile
        275                 280                 285

Leu Val Ser Asn Asp Trp Leu Phe Gly Phe Ser Ser Tyr Val Thr Asn
    290                 295                 300

Ile Met Asp Val Met Asp Arg Val Asn Pro Asp Gly Met Ala Phe Ile
305                 310                 315                 320

Pro Leu Arg Val Ile Pro Phe Leu Arg Glu Lys Gly Val Pro Gln Glu
                325                 330                 335

Thr Leu Ala Gly Ile Thr Val Thr Asn Pro Ala Arg Phe Leu Ser Pro
            340                 345                 350

Thr Leu Arg Ala Ser
        355
```